四种栎属青冈亚属植物叶绿体基因组特征及系统发育研究

黄婷，唐梦，陈晓丽，李卜宇，张雪梅*
（西华师范大学 生命科学学院，四川 南充 637002）

摘要：青冈亚属植物的系统发育地位长期存在着争议，部分种的种间关系不明确。为揭示宁冈青冈（*Quercus ningangensis*）、曼青冈（*Q. oxyodon*）、毛曼青冈（*Q. gambleana*）、竹叶青冈（*Q. neglecta*）的叶绿体基因组特征及系统发育关系，本研究选择以上 4 种青冈的成熟叶片进行二代测序，对其叶绿体基因组结构和特征进行分析，并结合相关类群进行系统发育研究。结果表明：（1）宁冈青冈、曼青冈、毛曼青冈、竹叶青冈的叶绿体基因组序列长度分别为：160 906 bp、160 883 bp、160 832 bp、160 784 bp，它们均编码 133 个基因，包括 88 个蛋白质编码基因、37 个 tRNA 基因、8 个 rRNA 基因。（2）四种青冈更偏好以 A/T 结尾的密码子，质体基因组变异区域主要存在于非编码区域。（3）通过 IR 边界分析得出，四种青冈亚属植物存在 ycf1 假基因，且在 IRb/SSC 区域发生扩张。（4）系统发育分析表明，在壳斗科中，水青冈属（*Fagus*）和三棱栎属（*Trigonobalanus*）较早分化出来，栎亚属（subg. *Quercus*）未形成一个单系群，叶绿体基因组建树结果和核标记、质体标记一致，栎亚属中 Ilex 组和 Cerris 组间出现穿插现象。（5）《中国植物志》及部分地方植物志将毛曼青冈作为一个单独的种处理，而周浙昆、邓敏等人根据叶的特征将毛曼青冈作为曼青冈的一个亚种处理，毛曼青冈的系统发育地位存在着争议,本实验根据叶绿体基因组信息并结合前人的形态分析结果，支持毛曼青冈作为一个独立的种存在的观点。该研究为探讨青冈亚属系统发育地位、青冈亚属组的划分、存疑种种间关系的解决提供了基础资料。

关键词：壳斗科，青冈亚属，基因组比较，叶绿体基因组，系统发育研究

Comparison of chloroplast genomes and phylogenic analysis of four Species of *Quercus* subg. *Cyclobalanopsis*

HUANG Ting, TANG Meng, CHEN Xiaoli, LI Puyu, ZHNAG Xuemei*
(*College of Life Sciences*, *China West Normal University*, Nanchong 637002, Sichuan, China)

Abstract: The phylogenetic status of *Quercus* subg. *Cyclobalanopsis* has long been controversial and the interspecific relationships of some species are unclear. At present, chloroplast genomes have been used to solve the phylogenetic problems of subg. *Quercus* and *Castanea* in the Fagaceae, but there are few researches on solving the phylogenetic problems of *Quercus* subg.

基金项目：教育部春晖计划项目 (467305)；全国第四次中药普查项目 (2018CP013)；国家标本平台教学标本子平台 (http://mnh.scu.edu.cn/) (2005DKA21403-JK)；西华师范大学科研创新团队资金资助 (KCXTD2022-4）。[Supported by Chunhui Project of Ministry of Education (467305); The Fourth National TCM General Survey Project (2018CP013); National Specimen Platform Teaching Specimen Sub-platform (http://mnh.scu.edu.cn/) (2005DKA21403-JK); Supported by the Innovation Team Fund of China West Normal University (KCXTD2022-4)]。

第一作者：黄婷(1997-)，硕士研究生，主要从事木本植物系统发育研究，（E-mail）ht929170315@163.com。
***通信作者：**张雪梅，博士，副教授，主要从事植物系统与进化研究，（E-mail）zhangmei103127@sina.com。

Cyclobalanopsis with chloroplast genome information. To reveal the chloroplasts genome characteristics and phylogenetic relationships of *Q. ningangensis*, *Q. oxyodon*, *Q. gambleana* and *Quercus neglecta*. In this study, mature leaves of the above four species of *Quercus* subg. *Cyclobalanopsis* were selected for Next-generation sequencing, and the chloroplast genome structure and characteristics were analyzed, Phylogeny was studied in combination with related taxa. The results were as follows: (1) The chloroplast genome sequences of *Q. ningangensis*, *Q. oxyodon*, *Q. gambleana* and *Q. neglecta* were 160 906 bp, 160 883 bp, 160 832 bp and 160 784 bp, respectively. They all encoded 133 genes, including 88 protein-coding genes, 37 tRNA genes and 8 rRNA genes. (2) The codons ending in A/T were preferred by the four species of *Quercus* subg. *Cyclobalanopsis*, and the variation regions of plastid genome mainly existed in non-coding regions. (3) According to IR boundary analysis, *ycf1* pseudogene was found in four species of *Quercus* subg. *Cyclobalanopsis* and expanded in IRb/SSC regions. (4) Phylogenetic analysis showed that *Fagus* and *Trigonobalanus* were differentiated early in Fagaceae, while subg *Quercus* did not form a monophyly. The results of phylogenetic tree based on chloroplast genome were consistent with the nuclear marker plasmid markers. Interspersed between Ilex group and Cerris group in subg *Quercus*. (5) The Flora of China and some local flora treated *Q. gambleana* as a separate species, while Zhou Zhekun and Deng Min et al. treated *Q. gambleana* as a subspecies according to the characteristics of leaves. The phylogenetic status of *Q. gambleana* is still controversial. Based on the chloroplast genome information and previous morphological analysis results, the present study supported the idea that *Q. gambleana* existed as an independent species. This study provides basic data for discussing the phylogenetic status of the *Quercus* subg. *Cyclobalanopsis*, the division of its groups, and the resolution of doubtful relationships among them.

Keywords: Fagaceae, *Quercus* subg. *Cyclobalanopsis*, genome comparison, chloroplast genome, phylogenic analysis

叶绿体广泛存于高等植物及大部分藻类中，是绿色植物进行光合作用的场所。叶绿体自身拥有独立的遗传物质，其基因组为四分体结构，由一个大单拷贝区域（long single copy, LSC）、一个小单拷贝区域（small simple copy, SSC）及两个反向重复区域（inverted repeat, IRA and IRB）构成，基因组序列和结构都相对保守（Zhao et al., 2018）。与核基因组比较结果显示，叶绿体基因组具有结构简单、分子量小、基因的替换率较低、单拷贝等特点（Korpelainen, 2004; Kwak et al., 2019）。近年，随着二代测序时间和成本的降低，越来越多的叶绿体基因组的测序分析结果已经用于模糊种的鉴定和系统发育研究。目前，已有研究利用叶绿体基因组解决壳斗科栎亚属（subg. *Quercus*）、栗属（*Castanea*）的系统发育问题（杨颜慈, 2018；高潇潇, 2020），但利用叶绿体基因组信息解决青冈亚属系统发育问题的还未多见。

壳斗科（Fagaceae）栎属青冈亚属（*Quercus* subg. *Cyclobalanopsis*）共有 122 种（Frodin & Govaerts, 1998）或 150 种（Huang et al., 1999），主要分布于亚洲热带和亚热带地区。我国是青冈亚属的重要分布区和多样性中心之一（罗艳和周浙昆, 2001a），共有 77 种以及 3 变种，秦岭、淮河以南为青冈亚属植物的主要分布区，为常绿阔叶林的建群种之一（陈焕镛和黄成就, 1998；Huang et al., 1999）。青冈类植物木材具有质地坚硬、耐腐蚀、耐磨损等优点，是制作船桨和其他工具的良好木材；其果实内淀粉含量丰富，含量约为 50%~60%（端木炘, 1995），可用于酿酒、制作饲料、糊料和工业淀粉等；树皮中含有单宁，可制作栲胶。青冈类植物的系统发育地位长期存在着争议，部分学者根据壳斗的排列方式将青冈类植物作为一个单独的属处理（徐永椿和任宪威, 1976；郑万钧, 1985；陈焕镛和黄成就, 1998；Huang et al., 1999）；但多数学者根据花粉粒解剖学等形态特征及 ITS 序列，支持青冈类植物作为

栎属下的一个亚属处理（中国科学院植物研究所, 1972；刘兰芳和房志坚, 1986；端木炘, 1992；罗艳和周浙昆, 2001b）。由于青冈类植物生殖结构变化较大，叶的特征较为均一，不同类群间存在基因渐渗，馆藏的标本量有限，果实等重要分类特征难以观察（邓敏, 2007），故其属下没有明确的分组，部分种的种间关系不明确，种的划分上还存在着许多争议。目前关于青冈亚属系统发育的研究主要还停留在形态学方面。因此，需要更多的分子系统学方面的证据来协助解决青冈亚属的系统发育地位和种间关系。

本研究对 4 种青冈亚属植物的成熟叶片进行二代测序，依托 4 种青冈亚属植物的叶绿体基因组信息，拟探究以下问题：（1）4 种青冈亚属植物的叶绿体基因组基本特征及其差异。（2）4 种青冈亚属植物的密码子使用偏好性。（3）毛曼青冈和曼青冈亲缘关系如何？以期为青冈亚属系统发育问题解决、种间关系的划分、种的鉴定及优良青冈木材的选育从叶绿体基因组层面提供了新的资料。

1 材料与方法

1.1 实验材料

在野外采集宁冈青冈（*Quercus ningangensis*）、曼青冈（*Q. oxyodon*）、毛曼青冈（*Q. gambleana*）、竹叶青冈（*Q. neglecta*）新鲜的、无病虫害的成熟叶片（表 1），用湿润的纱布擦拭后置于装有变色硅胶的密封袋中保存，干燥后置于-80℃的冰箱内备用。凭证标本存放于西华师范大学标本馆（Herbarium, College of Life Sciences, China West Normal University）。

表 1 4 种青冈亚属植物样品采集信息

Table 1 Information on the collection of samples from four species of *Quercus* subg. *Cyclobalanopsis*

植物名 Plant name	采集地点 Collection place	经纬度 Latitude/Longitude	海拔(m) Altitude (m)	凭证标本号 Voucher specimen
宁冈青冈 *Quercus ningangensis*	湖南绥宁县 Suining, Hunan	N26°22′24.74″ E110°07′54.40″	123 3	HT211016
曼青冈 *Q. oxyodon*	广西兴安县猫儿山 Maoer Mountain, Xingan, Guangxi	N25°54′21.67″ E110°27′58.79″	149 8	HT210925
毛曼青冈 *Q. gambleana*	重庆市南川区 Nanchuan, Chongqing	N29°00′58.46″ E107°11′18.90″	771	HT211014
竹叶青冈 *Q. neglecta*	贵州铜仁市梵净山 Fanjin Mountain, Tongren, Guizhou	N27°53′48.29″ E108°43′19.03″	156 3	HT210917

1.2 基因组 DNA 的提取与测序

用改良的 CTAB 法从 4 种青冈亚属植物叶片中提取总 DNA。利用 Illumina Hiseq 对提取的 DNA 进行双端测序。叶绿体全基因组 DNA 的提取与测序均由南京奥维森基因科技有限公司完成。

1.3 组装与注释

利用 FASTQ 软件进行数据质控，将质控得到的 clean data 利用 GetOrganelle 软件（Jin et al., 2020）进行组装，将原始数据 map 到组装好的序列上查看覆盖率并结合 Gapcloser 软件

（Zuo et al., 2017），补齐组装出现的空缺（gaps）。在 NCBI （https://www.ncbi.nlm.nih.gov/）上下载褐叶青冈（MN199023)的叶绿体基因组作为参考基因组，利用 CPGAVAS2（Shi et al., 2019）（http://www.herbalgenomics.org/cpgavas/）进行注释，注释好的数据利用 Geneious 软件（Kearse et al., 2012）进行手动调整和修改。利用 OGDRAW 软件（https://chlorobox.mpimp-golm.mpg.de/OGDraw.html）绘制 4 种青冈亚属植物的叶绿体基因组物理结构图谱。最后将注释好的数据提交到 NCBI，宁冈青冈、曼青冈、毛曼青冈、竹叶青冈的序列登录号分别为：ON303301、ON258628、ON258629、ON258631。

1.4 叶绿体基因组比较分析

利用 Geneious（Kearse et al., 2012）及 Editseq 软件统计 4 种青冈亚属植物的叶绿体基因组序列 SSC、LSC、IRS 区域的长度、GC 含量等基本信息。从 NCBI 上下载小叶青冈（MN199025）、褐叶青冈（MN199023）、西畴青冈（NC_036941）、青冈（NC_036930）的叶绿体基因组与我们测得四种青冈亚属植物的叶绿体基因组进行比较分析。

选择以 ATG 为起始密码子、长度大于 300 pb 的非重复序列，所有物种筛选后均剩余 52 条 CDS 序列。利用 Codon W 软件（Sharp & Li, 1987）计算 4 种青冈亚属植物及其近缘种的同义密码子相对使用度（relative synonymous codon usage, RSCU）、密码子适应指数（codon adaptation index, CAI）、有效密码子数（effective number of codons, ENC）、密码子偏好性指数（codon bias index, CBI）、最优密码子使用频率（frequency of optimal codons, FOP）以及每条 CDS 序列的 ENC 值、RSCU 值及密码子第三位各碱基出现的概率（记为 A3、T3、C3、G3），运用在线软件 EMBOSS (https://www.bioinformatics.nl/emboss-explorer/) 计算各 CDS 序列 GC1、GC2、GC3 的含量。中性绘图分析以 GC1 和 GC2 的平均值（GC_{12}）为纵坐标，以 GC3 值为横坐标，绘制散点图，并插入 y=x 的函数图像。ENC-plot 分析绘制的散点图将 GC3 值、ENC 值分别作为横、纵坐标，添加理论 ENC 值曲线，标准曲线公式为 $ENC=2+GC3+29/[GC3^2+(1-GC3)^2]$。PR2-plot 分析以 G3/(G3+C3)为横坐标，A3/(A3+T3)为纵坐标绘制散点图，分析 A、T、C、G 的使用频率和偏倚性。中性绘图、ENC-plot 绘图、PR2-plot 绘图均使用 origin 软件。

利用 mVISTA 软件（https://genome.lbl.gov/vista/mvista/instructions.shtml）（Mayor et al., 2000）对 4 种青冈亚属植物及其近缘种的叶绿体基因组进行可视化对比分析。使用 IRscope（https://irscope.shinyapps.io/irapp/）（Amiryousefi et al., 2018）绘制宁冈青冈、曼青冈、毛曼青冈、竹叶青冈及其近缘种的叶绿体基因组 IR 边界。

1.5 系统发育分析

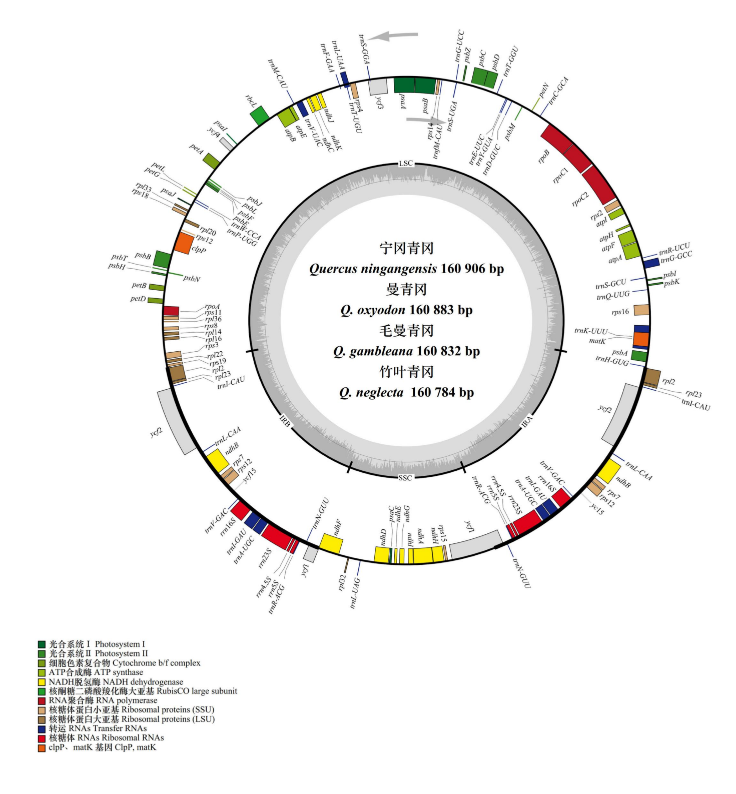
选取并下载在 NCBI 中公开释放的 8 种壳斗科植物的叶绿体基因组序列，包括：倒卵叶青冈（*Q. obovatifolia*，MG356785.1）、川滇高山栎（*Q. aquifolioides*，KX911971.1）、刺叶高山栎（*Q. spinosa*，MG678038.1）、巴东栎（*Q. engleriana*，MZ196209.1）、栓皮栎（*Q. variabilis*，NC031356.1）、三棱栎（*Trigonobalanus doichangensis*，NC023959.1）、台湾水青冈、米心水青冈选用台湾水青冈（*Fagus hayatae*）和米心水青冈（*F. engleriana*）作为外类群，将上述 8 种壳斗科植物的叶绿体基因组与本文组装的 4 种青冈亚属植物的叶绿体基因组一起构建系统发育树。将所有叶绿体基因组序列运用 MAFFT（Katoh & Standley, 2013）进行对比，将对比好的叶绿体基因组序列利用 MEGA 7 软件（Kumar et al., 2016）对序列开头和结尾未对齐部分进行适当的修剪，运用 MEGA 7 软件中 Models 功能寻找以上序列运用 ML 法构建系统发育树的最佳模型（GTR+GI），以最大似然法（maximum likelihood，ML），设置 bootstrap 为 100 0 构建系统发育树。另将修剪好的序列再运用 MEGA 7 软件以邻接法（neighbor-joining method, NJ），选用 Maximum Composite Likelihood 模型，设置 bootstrap 为 100 00 构建相关物种的系统发育树。

2 结果与分析

2.1 叶绿体基因组的基本特性

宁冈青冈、曼青冈、毛曼青冈、竹叶青冈的叶绿体基因组均有由一个 LSC 区域、两个 IRs 区域、一个 SSC 区域构成，为典型的四分体结构（图 1）。其基因组长度为 160 681~160 906 bp，其中曼青冈和毛曼青冈叶绿体全基因组的长度分别为 160 883 bp 和 160 882 bp，仅相差 1 个碱基。LSC 区的长度范围为 90 245~90 360 bp，SSC 区的长度范围为 18 891~18 929 bp，IRs 的长度范围为 25 816~25 840 bp。本实验测得的四种青冈的总 GC 含量（36.9%）、IRs 区域的 GC 含量（42.8%）及 CDs 区域的 GC 含量（37.9%）均一致，LSC 区域及 SSC 区域的 GC 含量也相差不大（表 2）。

宁冈青冈、曼青冈、毛曼青冈、竹叶青冈的叶绿体基因组注释结果显示，4 种青冈亚属植物均有 133 个基因，其中包含 37 个 tRNA 基因（transfer RNA genes），8 个 rRNA 基因（ribosomal RNA genes），以及 88 个蛋白质编码基因（Protein-coding Genes），（表 2）。有 15 个基因有 1 个内含子（*rpoC1*、*ndhA*、*ndhB*、*rpl2*、*rpl16*、*atpF*、*rps16*、*trnL-UAA*、*trnK-UUU*、*trnI-GAU*、*trnA-UGC*、*trnG-GCC*、*trnV-UAC*、*petB*、*petD*），包含 2 个内含子的基因有 3 个（*rps12*、*clpP*、*ycf3*）（表 3）。



圆圈内外的灰色箭头表示基因转录方向；不同的颜色代表不同功能的基因；内圈深灰色表示 GC 含量，浅灰色表示 AT 含量。

Gray arrows inside and outside the circle indicate the direction of gene transcription; Different colors represent genes with different functions; Dark gray in the inner circle represents GC content, the light gray represents AT content.

图 1 4 种青冈亚属植物的叶绿体基因组图谱

Fig. 1 Gene map of chloroplast genome in four species of *Quercus* subg. *Cyclobalanopsis*

表 2 4 种青冈亚属植物叶绿体基因组特征比较

Table 2 Comparison of chloroplast genomes among four species of *Quercus* subg. *Cyclobalanopsis*

表 3 4 种青冈亚属植物叶绿体基因组基因组成

物种 Species	宁冈青冈 *Quercus ningangensis*	曼青冈 *Q. oxyodon*	毛曼青冈 *Q. gambleana*	竹叶青冈 *Q. neglecta*
基因组大小 Genome size (bp)	160 906	160 883	160 832	160 784
大单拷贝区长度 Length of LSC (bp)	90 297	90 360	90 259	90 245
小单拷贝区长度 Length of SSC (bp)	18 929	18 891	18 905	18 895
反向重复区长度 Lengh of IRs (bp)	25 840	25 816	25 834	25 822
蛋白编码基因数 Number of protein-coding genes	88	88	88	88
转运 RNA 基因数 Number of tRNA genes	37	37	37	37
核糖体 RNA 基因数 Number of rRNA genes	8	8	8	8
总基因数 Total number of genes	133	133	133	133
总的 GC 含量 GC content of overall (%)	36.9	36.9	36.9	36.9
大单拷贝区 GC 含量 GC content of LSC (%)	34.8	34.8	34.7	34.8
小单拷贝区 GC 含量 GC content of SSC (%)	31.1	31.1	31.1	31.0
反向重复区 GC 含量 GC content of IRs (%)	42.8	42.8	42.8	42.8
基因编码区 GC 含量 GC content of CDs (%)	37.9	37.9	37.9	37.9

Table 3 Composition of complete chloroplast genome of four species of *Quercus* subg. *Cyclobalanopsis*

基因功能分类 Category of gene	基因分组 Group of gene	基因名称 Name of gene	数量 Number
自我复制 Self-replication	核糖体 RNA 基因基 rRNA genes	*rrn4.5S(×2); rrn5S(×2); rrn16S(×2); rrn23S(×2)*	8

	转运 RNA 基因 tRNA genes	*trnA-UGC(×2)*; trnC-GCA; trnD-GUC; trnE-UUC; trnF-GAA; trnfM-CAU; trnG-GCC*; trnG-UCC; trnH-GUG; trnI-CAU(×2); trnI-GAU(×2)*; trnK-UUU*; trnL-CAA(×2); trnL-UAA*; trnL-UAG; trnM-CAU; trnN-GUU(×2); trnP-UGG; trnQ-UUG; trnR-ACG(×2); trnR-UCU; trnS-GCU;trnS-GGA; trnS-UGA;trnT-GGU; trnT-UGU; trnV-GAC(×2); trnV-UAC*; trnW-CCA; trnY-GUA*	37
	RNA 聚合酶 RNA polymerase	*rpoA; rpoB; rpoC1*; rpoC2*	4
	核糖体小亚基 Small subunit of ribosomal	*rps2; rps3; rps4; rps7(×2); rps8; rps11; rps12(×2)**; rps14; rps15; rps16*; rps18; rps19*	14
	核糖体大亚基 Large subunit of ribosomal	*rpl2(×2)*; rpl14; rpl16*; rpl20; rpl22; rpl23(×2); rpl32; rpl33; rpl36*	11
	NADH-脱氢酶 NADH dehydrogenase	*ndhA*; ndhB(×2)*; ndhC; ndhD; ndhE; ndhF; ndhG; ndhH; ndhI; ndhJ; ndhK*	12
光合作用 Photosynthesis	光系统 1 Photosystem I	*psaA; psaB; psaC; psaI; psaJ*	5
	光系统 2 Photosystem II	*psbA; psbB; psbC; psbD; psbE; psbF; psbH; psbI; psbJ; psbK; psbL; psbM; psbN; psbT; psbZ*	15
	细胞色素 b/f 复合物	*petA; petB*; petD*; petG; petL;*	6

	Cytochrome b/f complex	*petN*	
	ATP 合成酶 ATP synthase	*atpA; atpB; atpE; atpF*; atpH; atpI*	6
	二磷酸核酮糖羧化酶大亚基基因 Large subunit of rubisco	*rbcL*	1
其他基因 Other genes	蛋白酶基因 Protease	*clpP***	1
	成熟酶基因 Maturase	*matK*	1
	乙酰 CoA 羧化酶的亚基 Subunits of Acetyl-CoA-carboxylase	*accD*	1
	包膜蛋白基因 Envelope membrane protein	*cemA*	1
	C 型细胞色素合成 C-type cytochromesynthesis	*ccsA*	1
	翻译起始因子 Translation initiation factor	*infA*	1
未知功能基因 Unknown function gene	开放阅读框 Open reading frames	*ycf1(×2); ycf2(×2); ycf3**; ycf4; ycf15(×2)*	8
总数 Total			133

注：（×2）两次拷贝；*一个内含子；**两个内含子。

Note: (×2) two copies; * One intron; ** Two introns.

2.2 密码子偏好性分析

2.2.1 密码子组成分析

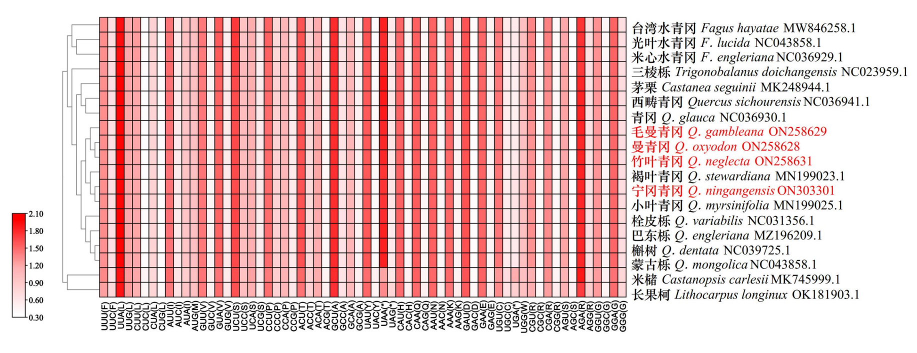
4 种青冈亚属植物的密码子数在 20 996~20 977 之间；有效密码子数（effective number of codons, ENC）在 49.91~49.81 之间，密码子适应指数（codon adaptation index, CAI）均为 0.167；宁冈青冈、毛曼青冈和竹叶青冈的密码子偏好性指数（codon bias index, CBI）为-0.099，曼青冈的密码子偏好性指数为-0.097；4 种青冈亚属植物的最优密码子使用频率（frequency of optimal codons, FOP）均为 0.355，GC 含量为 37.93%~37.95%之间。4 种青冈亚属植物的密码子偏好性各指数相差不大，说明它们密码子使用情况相似（表 4）。4 种青冈亚属植物的 GC3 含量在 29.85%~29.88%之间，表明青冈亚属植物更偏好以 A/T 结尾的密码子。

如图二，亮氨酸（Leu）、丝氨酸（Ser）精氨酸（Arg）有 6 个同义密码子，色氨酸（Trp）、甲硫氨酸（Met）仅由一种密码子编码，其余氨基酸均由 2 个及以上的同义密码子编码。这 8 种青冈亚属植物的密码子偏好性相似，其中 ACU、UCU、UUA、GCU、UAU、GAU、AGA 的使用频率较高，CUC、CUG、GCG、UAC、CAC、CAG、AAC、AAG、GAC、GAG、CGC、CGG、AGC、GGC 的使用频率较低。在米槠（*Castanopsis carlesii*）、长果柯（*Lithocarpus longinux*）中 UAA 的使用频率较所列的其他植物低。

表 4 4种青冈亚属植物的密码子偏好性指数

Table 4 Codon preference index of four species of *Quercus* subg. *Cyclobalanopsis*

指数 Index	宁冈青冈 *Quercus ningangensis*	曼青冈 *Q. oxyodon*	毛曼青冈 *Q. gambleana*	竹叶青冈 *Q. neglecta*
长度 Length (bp)	63 021	62 982	62 991	63 039
密码子数 Codon number	20 990	20 977	20 980	20 996
有效密码子数 Effective number of codons	49.89	49.90	49.91	49.91
密码子适应指数 Codon adaptation index	0.167	0.167	0.167	0.167
密码子偏好指数 Codon bias index	-0.099	-0.097	-0.099	-0.099
最优密码子使用频率 Frequency of optimal codons	0.355	0.355	0.355	0.355
GC 含量 GC content (%)	37.93	37.95	37.94	37.95
GC1 含量 GC1 content (%)	46.12	46.12	46.11	46.12
GC2 含量 GC2 content (%)	37.84	37.87	37.85	37.85
GC3 含量 GC3 content (%)	29.85	29.87	29.87	29.88



注：白色白色和红色分别表示较低和较高的 RSCU 值。毛曼青冈、曼青冈、竹叶青冈、宁冈青冈为该研究组装和注释，其余物种均为 NCBI 上下载。

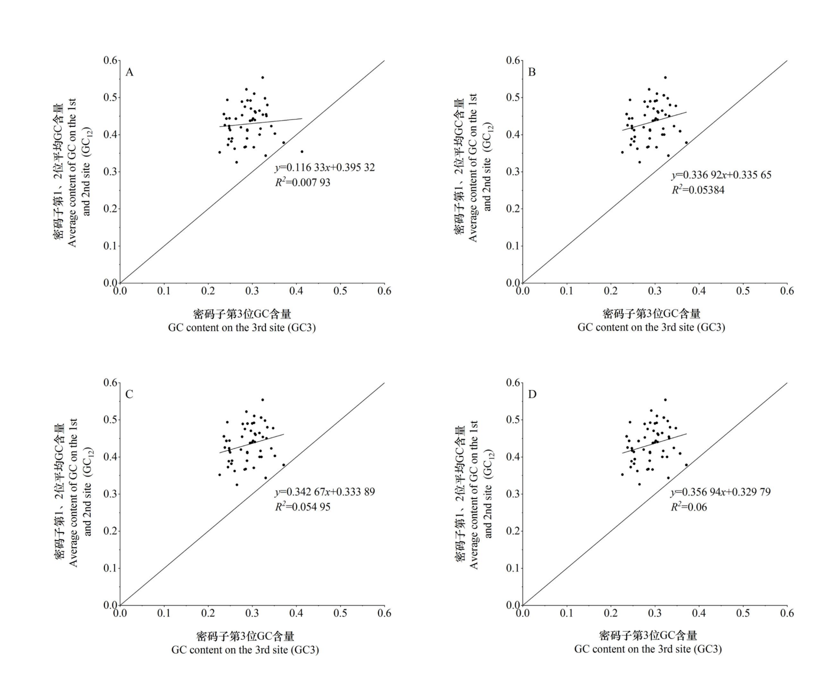
Note:White and red indicate lower and higher RSCU values, respectively. *Quercus gambleana*, *Q. oxyodon*, *Q. neglecta* and *Q. ningangensis* were assembled and annotated for the study, and the remaining species were downloaded from NCBI.

图 2 4种青冈亚属植物及近缘种的同义密码子的相对使用度

Fig. 2 RSCU values of all protein-coding (CDS) genes for chloroplast genomes from four species of *Quercus* subg. *Cyclobalanopsis* and its sibling species

2.2.2 中性绘图分析

从 4 种青冈亚属植物叶绿体基因组筛选出以 ATG 为起始密码子、长度大于 300 pb 的非重复序列进行中性绘图分析结果显示（图 3），宁冈青冈、曼青冈、毛曼青冈、竹叶青冈 GC3 和 GC_{12} 的相关系数分别为 0.007 93、0.053 84、0.005 495、0.06，回归系数分别为 0.116 33、0.336 92、0.342 67、0.356 94，双尾检验结果均为 P<0.01，显示出 GC_{12} 和 GC3 之间的相关性较弱，组成密码子碱基的第 1、2 位和第 3 位之间存在显著差异。推测自然选择对 4 种青冈亚属植物叶绿体基因组密码子的使用偏好性的影响更大。



A. 宁冈青冈叶绿体基因组中性绘图；**B.** 曼青冈叶绿体基因组中性绘图；**C.** 毛曼青冈叶绿体基因组中性绘图；**D.** 竹叶青冈叶绿体基因组中性绘图。

A. Neutrality polt analysis of the chloroplast genome of *Quercus ningangensis*; **B.** Neutrality polt analysis of the chloroplast genome of *Q. oxyodon*; **C.** Neutrality polt analysis of the chloroplast genome of *Q. gambleana*; **D.** Neutrality polt analysis of the chloroplast genome of *Q. neglecta*.

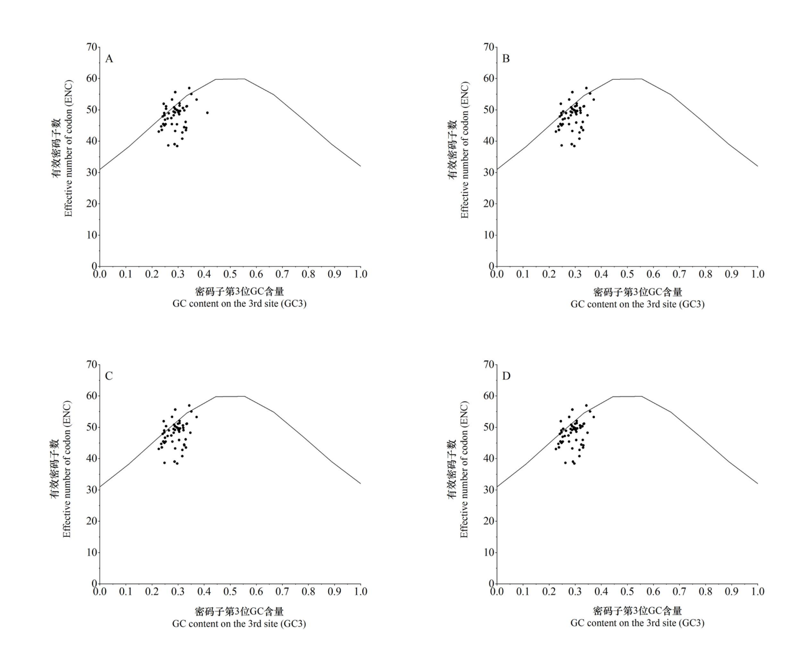
图 3 中性绘图分析

Fig. 3 Analysis of neutrality plot

2.2.3 ENC-plot 绘图分析

4 种青冈亚属植物的 ENC-plot 分析结果显示（图 4），大部分基因的实际 ENC 值在标准曲线之下，说明自然选择对这部分基因的密码子使用偏好性的影响更大；也有少部分基因的实际 ENC 值在标准曲线附近或上方，说明突变对其密码子的使用偏好性影响更大。ENC 比值频率分布（表 5）显示，宁冈青冈、曼青冈、毛曼青冈、竹叶青冈中位于-0.05~0.05 区间的基因分别有 19、18、18、19 个，分别占 37%、35%、35%、37%，表明突变对这些基因的影响较大。在这 4 种青冈亚属植物中还有 33~34 个基因的实际 ENC 值与理论 ENC 值差距较大，这部分基因受自然选择的影响作用较大。ENC-plot 分析结果和 ENC 比值频率分布

显示，4 种青冈亚属植物的叶绿体基因组密码子偏好性同时受突变和自然选择的影响，但总的来说，自然选择的影响大于突变带来的影响。



A. 宁冈青冈叶绿体基因组 ENC-plot 分析；**B.** 曼青冈叶绿体基因组 ENC-plot 分析；**C.** 毛曼青冈叶绿体基因组 ENC-plot 分析；**D.** 竹叶青冈叶绿体基因组 ENC-plot 分析。

A. ENC plot analysis of the chloroplast genome of *Quercus ningangensis*; **B.** ENC plot analysis of the chloroplast genome of *Q. oxyodon*; **C.** ENC plot analysis of the chloroplast genome of *Q. gambleana*; **D.** ENC plot analysis of the chloroplast genome of *Q. neglecta*.

图 4 ENC-plot 分析

Fig. 4 Analysis of ENC-plot

表 5 ENC 比值频数分布

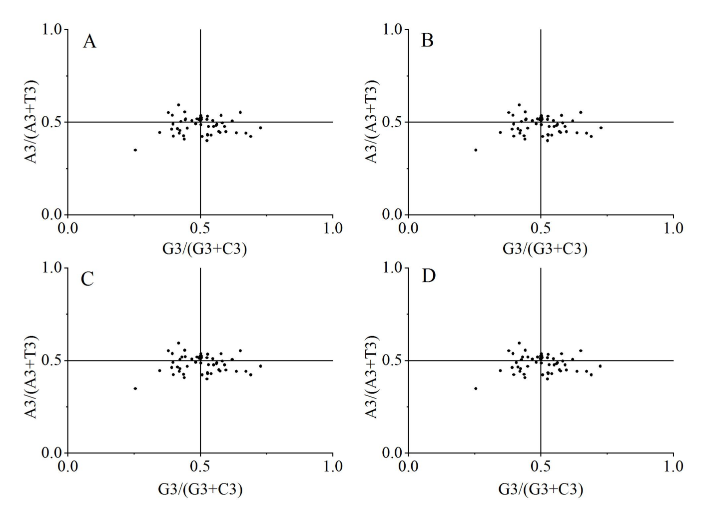
Fig. 4 Distribution of ENC ration

物种 Species	组限 Class boundary	组中值 Group mid-value	频数 Frequency number	组频 Frequency
	-0.15~-0.05	-0.08	3	0.06
宁冈青冈	-0.05~0.05	0.01	19	0.37
Quercus ningangensis	0.05~0.15	0.06	19	0.37
	0.15~0.25	0.19	10	0.19

	0.25~0.35	0.26	1	0.02
曼青冈 *Q. oxyodon*	-0.15~-0.05	-0.08	3	0.06
	-0.05~0.05	0.02	18	0.35
	0.05~0.15	0.06	21	0.40
	0.15~0.25	0.20	9	0.17
	0.25~0.35	0.26	1	0.02
毛曼青冈 *Q. gambleana*	-0.15~-0.05	-0.08	3	0.06
	-0.05~0.05	0.02	18	0.35
	0.05~0.15	0.07	21	0.40
	0.15~0.25	0.20	9	0.17
	0.25~0.35	0.26	1	0.02
竹叶青冈 *Q. neglecta*	-0.15~-0.05	-0.08	3	0.06
	-0.05~0.05	0.02	19	0.37
	0.05~0.15	0.07	20	0.38
	0.15~0.25	0.20	9	0.17
	0.25~0.35	0.26	1	0.02

2.2.4 PR2-plot 绘图分析

PR2-plot 绘图中（图 5），分布于中心点上或附近的基因主要受突变的影响，大多远离中心的点则受自然选择的影响。4 种青冈亚属植物的 PR2-plot 绘图分析显示，基因的分布不均匀，大部分基因分布于右下角，说明 4 种青冈亚属植物叶绿体基因组中密码子第 3 位碱基的使用频率为 A>T，G>C。表明 4 种青冈亚属植物密码子的偏好性受到了突变和自然选择的共同影响。



A3、T3、G3、C3 分别表示密码子第三位上各个碱基的含量。**A.** 宁冈青冈叶绿体基因组 PR2-plot 分析；**B.** 曼青冈叶绿体基因组 PR2-plot 分析；**C.** 毛曼青冈叶绿体基因组 PR2-plot 分析；**D.** 竹叶青冈叶绿体基因组 PR2-plot 分析。

A3, T3, G3, and C3 indicate the content of each base in the 3rd position of the codon. **A.** PR2-plot of chloroplast genes in *Quercus ningangensis*; **B.** PR2-plot of chloroplast genes in *Q. oxyodon*; **C.** PR2-plot of chloroplast genes in *Q. gambleana*; **D.** PR2-plot of chloroplast genes in *Q. neglecta*.

图 5 PR2-plot 分析

Fig. 5 Analysis of PR2-plot

2.3 叶绿体基因组差异比较分析

选宁冈青冈的叶绿体基因组做参考，利用 mVISTA 对曼青冈、毛曼青冈、竹叶青冈、小叶青冈、褐叶青冈、西畴青冈、青冈的叶绿体基因组序列进行可视化比较分析。从分析结果可以看出青冈亚属植物的叶绿体基因组具有较高的保守性，主要的变异区域位于非编码区，而 *ycf1* 基因在蛋白质编码区域发生变异较大（图 6）。LSC 区域的变异较大，IRs 区域变异小较为保守，SSC 区域的变异小于 LSC 区域。8 种青冈亚属植物的 rRNA 基因高度保守，几乎没有发生变异（图 6）。

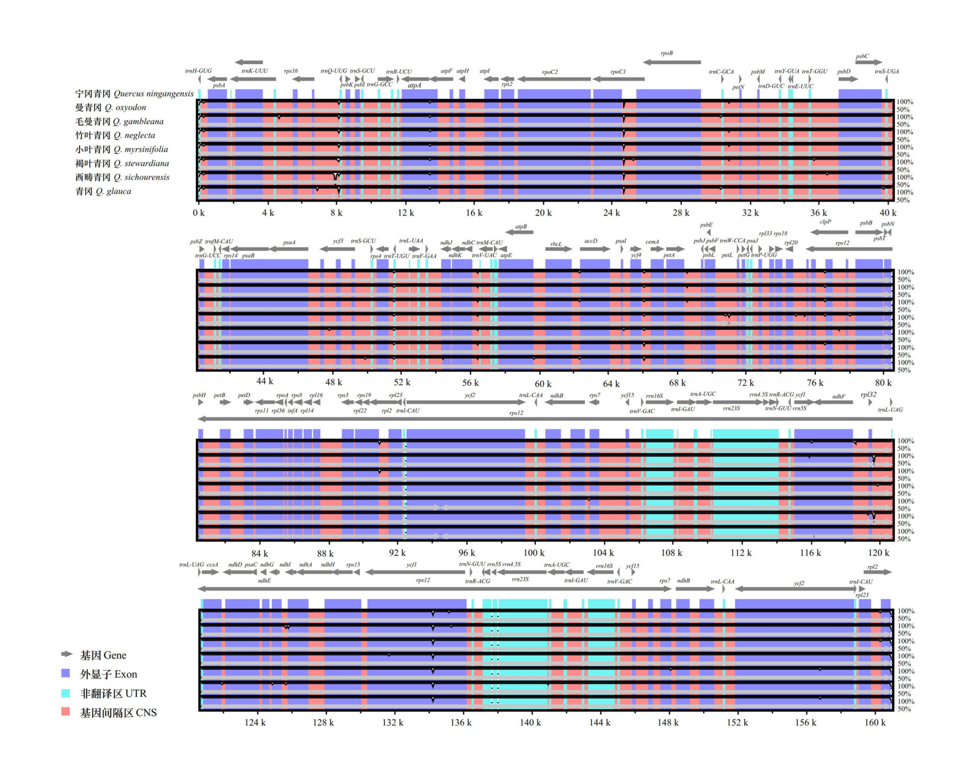
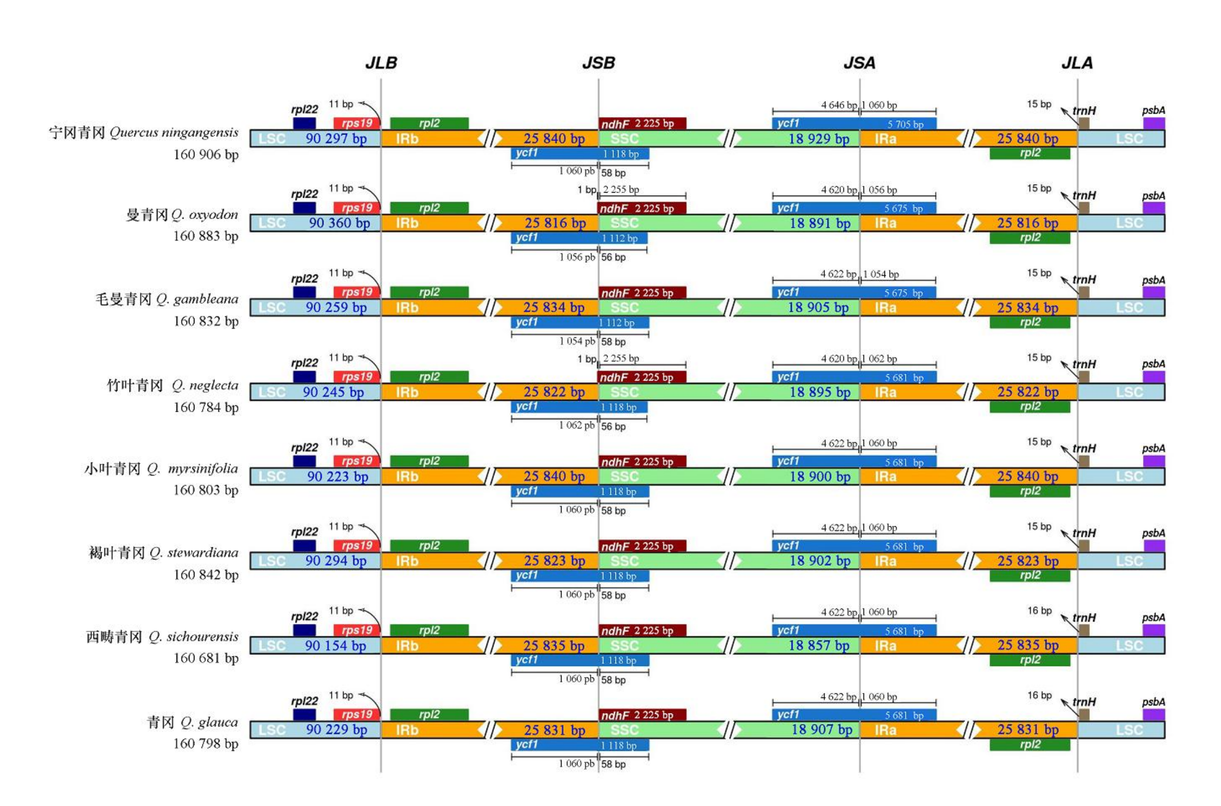


图 6 4 种青冈亚属植物及近缘种叶绿体基因组的可视化比对

Fig. 6 Visualization alignment of four chloroplast genomes in *Quercus* subg. *Cyclobalanopsis* and its related species

2.4 IR 边界分析

使用在线软件 IRscope 分析 4 种青冈亚属植物及其近缘种叶绿体基因组的 IR 扩张和收缩情况，以及识别 LSC、IR、SSC 的边界位置（图 7）。在 8 种青冈中，各边界附近分布的基因一致，但距边界的距离略有不同。其中，*rps19* 基因距 SSC/IRb 边界均为 11 bp，西畴青冈和青冈的 *trnH* 基因距离 IRa/LSC 边界的距离为 16 bp,其余青冈为 1 bp。在毛曼青冈和竹叶青冈中 *ndhF* 基因，有一个碱基位于 IRb 区域，其余 2 255 bp 均位于 SSC 区域，其余青冈中 *ndhF* 基因的长度为 2 255 bp，且均位于 SSC 区域。高等植物的叶绿体基因组中 *ycf1* 基因为第二长的基因，其横跨这 8 种青冈叶绿体基因组的 SSC/IRa 边界,在宁冈青冈、小叶青冈、褐叶青冈、西畴青冈、青冈中，*ycf1* 基因有 1 060 bp 位于 IRa 区域，在曼青冈、毛曼青冈、竹叶青冈中分别有 1 056 bp、1 054 bp、1 062 bp 位于 IRa 区域，在 8 种青冈的 IRb 区域也分别拷贝了相应长度的 *ycf1* 片段，这部分被视为 *ycf1* 假基因（ψ*ycf1*）。8 种青冈中，ψ*ycf1* 基因横跨 IRb/SSC，有 56~58 bp 位于 IRb/SSC 上游。



JLB. LSC/IRb 边界; **JSB**. IRb/SSC 边界; **JSA**. SSC/IRa 边界; **JLA**. IRa/LSC 边界。

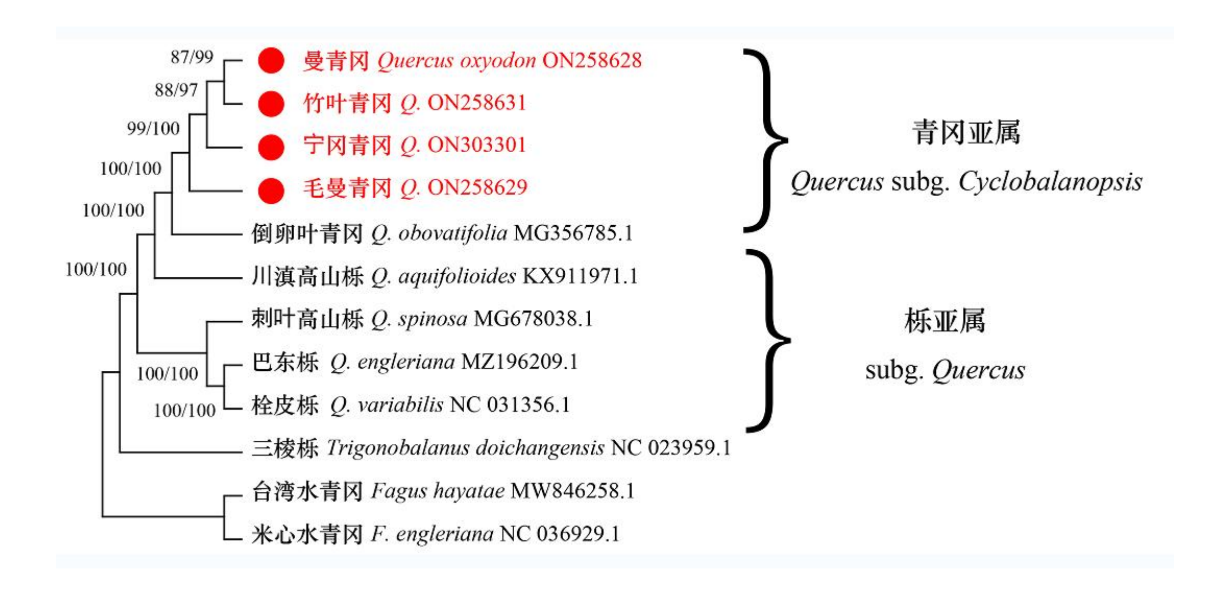
JLB. LSC/IRb boundary; **JSB**. IRb/SSC boundary; **JSA**. SSC/IRa boundary; **JLA**. IRa/LSC boundary.

图 7 4 种青冈亚属植物及近缘种的 LSC、IR 和 SSC 边界分析

Fig. 7 Comparison of LSC, SSC, and IR border regions among four chloroplast genomes in *Quercus* subg. *Cyclobalanopsis* and its sibling species

2.5 系统发育分析

选取 NCBI 上公开释放的 8 种壳斗科植物的叶绿体基因组与该研究组装的 4 种青冈亚属植物的叶绿体基因组序列，以台湾水青冈（*Fagus hayatae*）和米心水青冈（*F. engleriana*）作为外类群，分别用 ML 法和 NJ 法构建系统发育树，来探讨青冈亚属和栎亚属之间的关系及毛曼青冈和曼青冈的关系，并为青冈亚属系统发育问题的解决提供基础资料。两种方法构建的系统发育树均得到了较高的支持率且获得的拓扑结构一致。系统发育树显示（图 8），水青冈属（*Fagus*）和三棱栎属（*Formanodendron*）位于系统发育树的基部，为壳斗科中较早分化出来的类群。曼青冈和毛曼青冈成姐妹群关系，毛曼青冈为 4 种青冈亚属植物中最早分化出来的类群。栎亚属并未形成一个单系群，川滇高山栎（*Q. aquifolioides*）与青冈亚属的亲缘关系较近。



注：上面的节点数是支持值，左边是 ML 自展值，是 NJ 自展值。

Note: Number above nodes are support values with ML bootstrap values on the left, NJ bootstrap values on the right.

图 8 基于 ML 法和 NJ 法对 12 个物种叶绿体基因组序列构建系统发育树

Fig. 8 Phylogenetic relationships of the 12 species inferred from ML and NJ methods

3.讨论与结论

3.1 叶绿体基因组比较及密码子偏好性分析

被子植物的叶绿体全基因组长度在 120 000~170 000 bp 之间(Tangphatsornruang et al., 2010)，本研究中宁冈青冈、曼青冈、毛曼青冈、竹叶青冈叶绿体基因组的长度为 160 784~160 906 bp，在叶绿体基因组的大小、结构、基因的数量、总的 GC 含量及 4 个区域的长度和 GC 含量上都高度保守。青冈亚属的叶绿体基因组基本特征与其他壳斗科植物相似，壳斗科与其他科物种相比基因密度更大，对环境的适应更强，稳定性更高（高潇潇, 2020）。从 4 种青冈亚属植物的叶绿体全基因组比较分析可以看出，4 种青冈亚属植物的叶绿体基因组高度保守，变异区域主要在非编码区。青冈亚属植物的鉴定主要依据壳斗的特征，在没有果实的情况下，部分种难以鉴定，可以利用叶绿体基因组筛选变异较大的区域，开发成 DNA 条形码，便于青冈亚属植物的鉴定。通过对 8 种青冈亚属植物叶绿体基因组的研究，筛选出 *trnH-GUG~trnQ-UUG*、*ropC1~petN*、*trnT-UGC~trnM-CAU*、*ycf4~psbJ* 及 *ycf1* 区域作为 DNA 条形码，为后续青冈亚属植物的鉴定提供便利。被子植物叶绿体基因组 IR 区域的扩张和收缩普遍存在于植物进化过程中(Hansen et al., 2007; Davis & Soreng, 2010; Huang et al., 2014)。青冈亚属植物位于 IR 区域的基因相同，说明青冈亚属 IR 边界较为保守。4 种青冈亚属植物均出现 *ycf1* 假基因（ψ*ycf1*），在壳斗科栎亚属，栗属（*Castanea*）部分种中也出现了 *ycf1* 基因在 IRb/SSC 拷贝为假基因的现象（杨颜慈, 2018；高潇潇, 2020）。

密码子使用偏好性是生物进化的重要现象。主要受基因组的大小、碱基突变、基因漂变、自然选择、基因的表达水平、tRNA 丰度、GC 含量、蛋白质结构等因素影响（Romero et al., 2000；Duret , 2000；Angellotti et al., 2007）。从 GC3 的含量可以看出本研究中的青冈亚属植物更偏好以 A/T 结尾的密码子，这一现象在被子植物中广泛存在（Clegg et al., 1994；

Tangphatsornruang et al., 2010；Delannoy et al., 2011）。同义密码子相对使用度结果显示，青冈亚属和栎亚属植物的同义密码子使用度相似，米槠、长果柯中 UAA 的使用频率较所列的其他植物低，说明壳斗科植物的密码子使用偏好性相似，但不同的属之间仍存在着一定的差异。3 种绘图分析结果显示，4 种青冈亚属植物的密码子偏好性受选择压力和突变同时影响，但选择压力为主要影响因素，这与米槠（*Castanopsis carlesii*）（江淑珍等, 2021）、降香黄檀（*Dalbergia odorifera*）（原晓龙等, 2021）、灯盏花（*Erigeron breviscapus*）（李显煌等, 2021）等得出的结果相似。

3.2 系统发育关系及种间关系分析

系统发育树结果显示，水青冈属（*Fagus*）和三棱栎属（*Trigonobalanus*）分化较早，与前人根据化石和地理信息系统对壳斗科起源等研究的结果一致（周浙昆, 1999）。前人根据花粉特征及核基因将川滇高山栎、刺叶高山栎、巴东栎归为 Ilex 组、栓皮栎归为 Cerris 组（Hubert F et al., 2014），但核标记、质体标记的研究结果中显示，Ilex 组的亚洲种常出现穿插在 Cerris 组中的现象（Simeone MC et al., 2013；Hubert F et al., 2014），本研究中，4 种栎亚属植物并未形成一个单系群，刺叶高山栎（*Quercus spinosa*）与巴东栎（*Q. engleriana*）、栓皮栎（*Q. variabilis*）聚在一枝，Ilex 组和 Cerris 组之间出现穿插的现象，与核标记、质体标记的结果一致。在《中国植物志》（陈焕镛和黄成就, 1998）、《中国树木志》（郑万钧, 1985）、《云南植物志》（中国科学院昆明植物研究所, 1979）、《贵州植物志》（蓝开敏和李永康, 1982）、等地方植物志中，毛曼青冈和曼青冈是作为两个单独的种处理。罗艳和周浙昆（2001a）、邓敏（2007）等根据曼青冈、毛曼青冈叶背的星状毛等特征，认为毛曼青冈应作为曼青冈的一个变种处理。而刘凌燕等（2008）根据聚类分析的结果认为应将毛曼青冈作为长叶青冈（*Q. longifolia*）的变种处理。本研究中系统发育分析结果显示，曼青冈与竹叶青冈亲缘关系较近，曼青冈和毛曼青冈未形成姐妹群。毛曼青冈和曼青冈的果实和叶型相似，叶背毛的颜色和形态存在差异，但在不同地区采集的叶背毛差距很大（邓敏, 2007）。在扫描电镜下观察两者的花粉外壁纹饰纹饰上也存在区别，毛曼青冈的花粉外壁纹饰为颗粒疣状或聚合疣状，曼青冈的花粉外壁为颗粒状纹饰（王萍莉和溥发鼎, 2004）。本研究中采的毛曼青冈叶密被黄褐色星状毛，而曼青冈叶背的星状毛为白色且较毛曼青冈稀疏。综上，本研究更倾向于毛曼青冈和曼青冈作为两个种处理的观点。在本研究中毛曼青冈采自重庆，曼青冈采自广西，两地相距较远，我们推测可能由于生长环境存在较大差异及长期的基因交流受阻等因素导致本研究中的毛曼青冈和曼青冈逐渐形成两个种。

参考文献:

AMIRYOUSEFI A, HYVÖNEN J, POCZAI P, 2018. IRscope: an online program to visualize the junction sites of chloroplast genomes[J]. Bioinformatics, 34(17): 3030-3031.

ANGELLOTTI MC, BHUIYAN SB, CHEN GR et al., 2007. CodonO: codon usage bias analysis within and across genomes[J]. Nucl Acid Res, 35(suppl_2): W132-W136.

CHEN HY, HUANG CJ, 1998. Flora Reipublicae Popularis Sinicae: Tomus 22 [M]. Beijing: Science Press: 263-334. [陈焕镛, 黄成就, 1998. 中国植物志: 第二十二卷[M]. 北京: 科学出版社: 263-334]

CLEGG MT, GAUT BS, LEARN GH et al., 1994. Rates and patterns of chloroplast DNA evolution[J]. Proc Nat Acad Sci, 91(15): 6795-6801.

DAVIS JI, SORENG RJ, 2010. Migration of endpoints of two genes relative to boundaries between regions of the plastid genome in the grass family (Poaceae)[J]. Amer J Bot, 97(5):

874-892.
DELANNOY E, FUJII S, COLAS DES FRANCS-SMALL C et al., 2011. Rampant gene loss in the underground orchid Rhizanthella gardneri highlights evolutionary constraints on plastid genomes[J]. Molec Biol Evol, 28(7): 2077-2086.
DENG M, 2007. Anatomy, taxonomy, distribution & phylogeny of *Quercus* subg. *Cyclobalanopsis* (Oersted) Schneid (Fagaceae)[D]. Beijing: Chinese Academy of Sciences: 3-70. [邓敏, 2007. 壳斗科栎属青冈亚属的形态解剖、分类、分布及其系统演化[D]. 北京: 中国科学院研究生院: 3-70.]
DUAN MX, 1992. Studies on the Fagaceae wood-assortment in Taiwan[J]. Fujian Sci Technol, 19(4): 65-70. [端木炘, 1992. 台湾壳斗科木材分类的研究[J]. 福建林业科技, 19(4): 65-70.]
DUAN MX, 1995. The comprehensive utilization of the resources of *Cyclobalanopsis* in China[J]. J Beijing For Univ, 17(2): 109-110. [端木炘, 1995. 我国青冈属资源的综合利用[J]. 北京林业大学学报, 17(2): 109-110.]
DURET L, 2000. tRNA gene number and codon usage in the *C. elegans* genome are co-adapted for optimal translation of highly expressed genes[J]. Trends Genet, 16(7): 287-289.
FRODIN DG, GOVAERTS R, 1998. World Checklist and Bibliography of Fagales (Betulaceae, Corylaceae, Fagaceae and Ticodendraceae)[M]. London: the Royal Botanic Gardens: 201-394.
GAO XX, 2020. Comparative whole chloroplast genomics of *Castanea* and Fagales family[D]. Xian: Northwest UNIVERSITY: 14-29. [高潇潇, 2020. 壳斗目栗属植物比较叶绿体基因组与系统发育研究[D]. 西安: 西北大学: 14-29.]
HANSEN DR, DASTIDAR SG, CAI Z et al., 2007. Phylogenetic and evolutionary implications of complete chloroplast genome sequences of four early-diverging angiosperms: *Buxus* (Buxaceae), *Chloranthus* (Chloranthaceae), *Dioscorea* (Dioscoreaceae), and *Illicium* (Schisandraceae)[J]. Mol phylogenet evol, 45(2): 547-563.
HUANG CJ, ZHANG YT, BRUCE B, 1999. Flora of China Vol.4[M]. Beijing: Science Press and Missouri Botanical Garden Press: 314–400.
HUANG H, SHI C, LIU Y et al., 2014. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: genome structure and phylogenetic relationships[J]. BMC Evol biol, 14(1): 1-17.
HUBERT F, GRIMM GW, JOUSSELIN E et al., 2014. Multiple nuclear genes stabilize the phylogenetic backbone of the genus *Quercus*[J]. Syst Biodivers, 12(4): 405-423.
JIANG SZ, LIAN H, XIONG YF, et.al., 2021. Analysis of codon bias in chloroplast genome of *Castanopsis carlesii*[J]. Mol Plant Breed, 26(7): 1-12. [江淑珍, 连辉, 熊远芳等., 2021. 米槠叶绿体基因组密码子偏好性分析[J]. 分子植物育种, 26(7): 1-12.]
JIN JJ, YU WB, YANG JB et al., 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes [J]. Genome Biol, 21(1): 1-31.
KATOH K, STANDLEY DM, 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability[J]. Mol biol evol, 30(4): 772-780.
KEARSE M, MOIR R, WILSON A et al., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data[J]. Bioinformatics, 28(12): 1647-1649.
KORPELAINEN, 2004. The evolutionary processes of mitochondrial and chloroplast genomes differ from those of nuclear genomes [J]. Naturwissenschaften, 91(11): 505-518.
KUMAR S, STECHER G, TAMURA K, 2016. MEGA7: molecular evolutionary genetics analysis

version 7.0 for bigger datasets[J]. Mol biol evol, 33(7): 1870-1874.

KUNMING INSTITUTE OF BOTANY, CHINESE ACADEMY OF SCIENCES, 1979. Flora Yunnanica: Tomus 2 (spermatophyte)[M]. Beijing: Science Press: 322-325. [中国科学院昆明植物研究所, 1979. 云南植物志：第二卷（种子植物）[M]. 北京：科学出版社: 322-325]

KWAK SY, LEW TTS, SWEENEY CJ et al., 2019. Chloroplast-selective gene delivery and expression in planta using chitosan-complexed single-walled carbon nanotube carriers[J]. Nat nanotechnol, 14(5): 447-455.

LAN KM, LI YK, 1982. Guizhou botanical chronicles[M]. Guiyang: Guizhou People’S Press: 103-105. [蓝开敏，李永康, 1982. 贵州植物志[M]. 贵阳：贵州人民出版社: 103-105]

LIU LF, FANG ZJ, 1986. A study on pollen morphology of quercoideae oerst (fagaceae) in china[J]. Guihaia, 12(4): 243-251+323-327. [刘兰芳，房志坚, 1986. 中国壳斗科栎亚科花粉形态研究[J]. 广西植物, 12(4): 243-251+323-327.]

LIU LY, ZHANG ML, LI JQ, et.al., 2008. A Numerical Taxonomic Study of the Genus *Cyclobalanopsis* Oersted from China[J]. Wuhan Botan Res, 26(5): 466-475. [刘凌燕，张明理，李建强等., 2008. 国产青冈属的数量分类学研究[J]. 武汉植物学研究, 26(5): 466-475.]

LI XH, YANG SC, XIN YX, et.al., 2021. Analysis of the codon usage bias of chloroplast genome in *Erigeron breviscapus* (Vant.) Hand-Mazz[J]. J Yunnan Agr Univ(Nat Sci Ed), 36(3): 384-392. [李显煌，杨生超，辛雅萱等., 2021. 灯盏花叶绿体基因组密码子偏好性分析[J]. 云南农业大学学报(自然科学), 36(3): 384-392.]

LUO Y, ZHOU ZK, 2001a. Leaf epidermis of *Quercus* subgen. *Cyclobalanopsis* (Oerst.) Schneid. (Fagaceae)[J]. Acta Phytotaxon Sin, 39(6): 489-501. [罗艳，周浙昆, 2001a. 栎属青冈亚属(壳斗科)的叶表皮研究 [J]. 植物分类学报, 39(6): 489-501.]

LUO Y, ZHOU ZK, 2001b. Phytogeography of *Quercus* subg. *Cyclobalanopsis*[J]. Acta Bot Yunnanica, 23(1): 1-16+28. [罗艳，周浙昆, 2001b. 青冈亚属植物的地理分布[J]. 云南植物研究, 23(1): 1-16+28.]

MAYOR C, BRUDNO M, SCHWARTZ JR et al., 2000. VISTA: visualizing global DNA sequence alignments of arbitrary length[J]. Bioinformatics, 16(11): 1046-1047.

ROMERO H, ZAVALA A, MUSTO H, 2000. Codon usage in Chlamydia trachomatis is the result of strand-specific mutational biases and a complex pattern of selective forces[J]. Nucl acid res, 28(10): 2084-2090.

SHARP PM, LI WH, 1987. The codon adaptation index-a measure of directional synonymous codon usage bias, and its potential applications[J]. Nucl acid res, 15(3): 1281-1295.

SHI L, CHEN H, JIANG M et al., 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer[J]. Nucl acid Res, 47(W1): W65-W73.

SIMEONE MC, PIREDDA R, PAPINI A et al., 2013. Application of plastid and nuclear markers to DNA barcoding of Euro-Mediterranean oaks (Quercus, Fagaceae): problems, prospects and phylogenetic implications[J]. Bot J Linn Soc, 172(4): 478-499.

TANGPHATSORNRUANG S, SANGSRAKRU D, CHANPRASERT J et al., 2010. The chloroplast genome sequence of *mungbean (Vigna radiata) determined* by high-throughput pyrosequencing: structural organization and phylogenetic relationships[J]. DNA res, 17(1): 11-22.

WANG PL, PU DF, 2004. Pollen Morphology and Biogeography of Fagaceae[M]. Guangzhou: Guangdong Sci Technol Press: 1-158. [王萍莉，溥发鼎, 2004. 壳斗科植物花粉形态及生物地理[M]. 广州：广东科技出版社: 1-158]

XU YC, REN XW, 1976. The Classification and Distribution of Fagaceae of Yunnan Province (2)[J]. J Syst Evol, 14(2): 73-88. [徐永椿, 任宪威, 1976. 云南壳斗科分类与分布(2)[J]. 植物分类学报, 14(2): 73-88.]

YANG YC, 2018. Plastid genomes of *Quercus* in China and the key genera in Fagaceae: comparative and phylogenetic analyses[D]. Xi'an: Northwest UNIVERSITY: 19-90. [杨颜慈, 2018. 中国栎属植物和壳斗科主要属质体基因组比较分析和系统发育研究[D]. 西安: 西北大学: 19-90.]

YUAN XL, LI YQ, ZHANG JF, et.al., 2021. Analysis of codon usage bias in the chloroplast genome of *Dalbergia odorifera*[J]. Guihaia, 41(4): 622-630. [原晓龙, 李云琴, 张劲峰等., 2021. 降香黄檀叶绿体基因组密码子偏好性分析[J]. 广西植物, 41(4): 622-630.]

ZHAO Z, WANG X, YU Y et al., 2018. Complete chloroplast genome sequences of *Dioscorea*: Characterization, genomic resources, and phylogenetic analyses[J]. PeerJ, 6(12): e6032.

ZHENG WJ, 1985. Sylva Sinica: Vol. 2 [M]. Beijing: China Forestry Publishing House: 2199-2356. [郑万钧, 1985. 中国树木志 第 2 卷[M]. 北京: 中国林业出版社: 2199-2356]

ZHOU ZK, 1999. Fossils of the Fagaceae and their implications in systematics and biogeography[J]. Acta Phytotax Sin, 37(4): 369-385. [周浙昆, 1999. 壳斗科的地质历史及其系统学和植物地理学意义 [J]. 植物分类学报, 37(4): 369-385.]

ZUO LH, SHANG AQ, ZHANG S et al., 2017. The first complete chloroplast genome sequences of *Ulmus* species by de novo sequencing: Genome comparative and taxonomic position analysis[J]. PLoS ONE, 12(2): e0171264.